

Grenoble mardi 5 et mercredi 6 février 2019

Traitement intelligent de données industrielles, principes et exemples d'application

David HERNANDEZ

PhD, Ingénieur R&D – Automatique & Industrie

145 rue Louis Barran, 38430 Saint Jean de Moirans, France

david.hernandez@aifrance.com Tel : 04 76 93 79 90

Mots clés : *Traitement de données, outliers, reconstruction, méthode ensembliste, apprentissage semi-supervisé*

Résumé :

Nous présentons une méthodologie de traitement intelligent de données pour des applications industrielles. Cette méthodologie est conçue pour détecter des points aberrants (*outliers*) dans les variables d'une base de données et de proposer également une reconstruction aux observations détectés comme aberrantes.

La méthode proposée utilise une combinaison de techniques statistiques classiques avec des techniques issues de la data science et du « *machine learning* ». Nous avons séparé nos algorithmes en deux grands groupes de méthodes : les méthodes en univarié, c'est-à-dire en analysant uniquement la variable à nettoyer sans tenir compte de l'évolution des autres variables et les méthodes en multivarié, c'est-à-dire en prenant en compte l'évolution des autres variables par rapport au compteur cible à nettoyer. Le classificateur final est obtenu par une méthode ensembliste par moyenne des prédictions (régression). Nous utilisons notamment la classification des *outliers* par *clustering*, sous la prémisse d'un apprentissage semi-supervisé. Les résultats obtenus par nos algorithmes de détections sont très satisfaisants.

Nous décrivons le processus d'industrialisation de la solution et son implémentation sur deux cas d'étude. Deux exemples d'application réels sont présentés dans les domaines des plans de comptage pour l'efficacité énergétique industrielle ainsi que pour la maintenance prédictive sur une installation de production d'énergie renouvelable.